



**Europäisches  
Patentamt**

**European  
Patent Office**

**Office européen  
des brevets**

J1046 U.S. PRO  
09/843315  
04/25/01

**Bescheinigung**

**Certificate**

**Attestation**

Die angehefteten Unterla-  
gen stimmen mit der  
ursprünglich eingereichten  
Fassung der auf dem näch-  
sten Blatt bezeichneten  
europäischen Patentanmel-  
dung überein.

The attached documents  
are exact copies of the  
European patent application  
described on the following  
page, as originally filed.

Les documents fixés à  
cette attestation sont  
conformes à la version  
initialement déposée de  
la demande de brevet  
européen spécifiée à la  
page suivante.

**Patentanmeldung Nr.    Patent application No.    Demande de brevet n°**

00480035.5

Der Präsident des Europäischen Patentamts;  
Im Auftrag

For the President of the European Patent Office

Le Président de l'Office européen des brevets  
p.o.

**I.L.C. HATTEN-HECKMAN**

DEN HAAG, DEN  
THE HAGUE,  
LA HAYE, LE

09/06/00

**This Page Blank (uspto)**



Europäisches  
Patentamt

European  
Patent Office

Office européen  
des brevets

**Blatt 2 der Bescheinigung**  
**Sheet 2 of the certificate**  
**Page 2 de l'attestation**

Anmeldung Nr.:  
Application no.:  
Demande n°:

00480035.5

Anmeldetag:  
Date of filing:  
Date de dépôt:

26/04/00

Anmelder:  
Applicant(s):  
Demandeur(s):  
INTERNATIONAL BUSINESS MACHINES CORPORATION  
Armonk, NY 10504  
UNITED STATES OF AMERICA

Bezeichnung der Erfindung:  
Title of the invention:  
Titre de l'invention:

System and method for enabling best-effort service of subflow priority flows

In Anspruch genommene Priorität(en) / Priority(ies) claimed / Priorité(s) revendiquée(s)

Staat:  
State:  
Pays:

Tag:  
Date:  
Date:

Aktenzeichen:  
File no.  
Numéro de dépôt:

Internationale Patentklassifikation:  
International Patent classification:  
Classification internationale des brevets:

/

Am Anmeldetag benannte Vertragsstaaten:  
Contracting states designated at date of filing: AT/BE/CH/CY/DE/DK/ES/FI/FR/GB/GR/IE/IT/LI/LU/MC/NL/PT/SE  
Etats contractants désignés lors du dépôt:

Bemerkungen:  
Remarks:  
Remarques:

**This Page Blank (uspto)**

**SYSTEM AND METHOD FOR ENABLING  
BEST-EFFORT SERVICE  
OF SUBFLOW PRIORITY FLOWS**

**Field of the Invention**

5       The present invention relates generally to the quality of  
service (QoS) in communication networks, and more specifically  
applies to a system and method that warrant to priority flows a  
guaranteed traffic while subflows remain handled on a best  
effort basis so as they are possibly becoming eligible to  
10   discarding when lower priority flows are deemed to gain  
precedence.

## Background of the Invention

Network service providers and administrators of networks have long provided all their customers with the same level of performance namely, the best-effort service. However, in recent  
5 years, increased usage of private or public networks especially, the Internet has resulted in scarcity of network capacity, compromising performance of traditional, mission critical applications. At the same time, new time-critical applications have emerged which demand much improved service  
10 quality. For example, traffic for teleconferencing may require a higher priority to ensure smooth jitter-free audio and video. As a result, service providers have found mandatory to offer their customers alternative levels of service. To this end, Quality of Service (QoS) mechanisms, intended to provide to an  
15 application the necessary level of service (bandwidth and delay) in order to maintain an expected quality level, have been devised. While QoS may mean guaranteed bandwidth with zero frame loss to a mission-critical application it would rather mean guaranteed frame latency for a telephony application, such  
20 as voice over IP (VOIP). This level of fine-grain control that QoS would ideally provide, if not properly managed, would place a very significant burden on the network infrastructure since each device would have to keep an entry in its forwarding table for every flow it manages. Thus, in a large corporate network,  
25 devices would tend to become overwhelmed with millions of flows, especially at the boundaries. Although there is definitively a price to pay to cope with the explosion of traffic and to meet QoS requirements; protocols that actually deliver level of service by application flow such as ATM (Asynchronous Transport Mode), Frame Relay, and MPLS (Multi-Protocol Label Switch-  
30 ing) have rather implemented Class of Service (CoS) mechanisms instead, in an attempt to reduce flow complexity by mapping multiple flows into a few service levels. With this approach network resources are then allocated based on these service  
35 levels, and flows can be aggregated and forwarded according to

the service class of the packet. Instead of the fine grain control of QoS, not easily achievable, CoS applies bandwidth and delay to different classes of network services. CoS easily scales with network expansion. As the network grows, traffic  
5 continues to be managed based on a few service levels, keeping infrastructure burdens to a reasonable level.

A typical example of CoS is the recently ratified 802.1p IEEE standard for traffic prioritization at layer 2. Taking into account that different traffic classes have different  
10 priority needs this standard defines how network frames are tagged with user priority levels ranging from 7 (highest priority) to 0 (lowest priority). Therefore, 802.1p compliant network infra-structure devices, such as switches and routers, prioritize traffic delivery according to the user priority tag,  
15 giving higher priority frames precedence over lower priority or non-tagged frames. This means that time-critical data can receive preferential treatment over non-time-critical data.

This simplification is not however, without creating its own problems. An example to illustrate the difficulties encountered can easily be found when considering the distribution of  
20 video signals, a more frequent situation with the deployment of teleconferencing and the advent of video-on-demand services. Hence, digital video flows must be granted a high priority to avoid jitter in the receiving of the video frames so as the remotely reproduced animated pictures do not flicker, pause or  
25 just become blank pictures which is highly disturbing for those who are watching. However, video is also very demanding in term of bandwidth requirement (typically a few Mbits/sec per channel) and, because this type of traffic is also highly  
30 bursty, it is not infrequent that routing or switching devices, at nodes of telecommunication networks, recurrently experience congestion windows even though the network is on the overall well adapted to the traffic it must handle. When this happens, video flows, which have a high priority, prevent all the other  
35 ones from being served since they are easily able to preempt

all the available bandwidth, at a given output port of a node, for a sustainable period of time. This includes e.g., control messages exchanged between nodes, that do not deserve, under nominal conditions, to have a high priority and do not occupy much of the bandwidth however, playing a crucial role to insure that network be properly managed and indeed always runs at its optimal level of performances thus, contributing to prevent the kind of congestion window just described, to happen frequently. Now, although video signals must definitively be granted a high priority, it remains that digital coding of video signals is done in such a way that all the transmitted pieces are far to have the same importance for the device in charge of rebuilding images upon reception. MPEG which stands for Motion Picture Experts Group is the name of the ISO (International Standards Organization) committee that specified the compression, transmission and decompression scheme for video. The first objective of MPEG is to take analog or digital video signals and convert them to packets of digital information that are more efficiently transported over communication networks. Thus, MPEG is essentially aimed at compressing the video into much less information consuming less transmission bandwidth (only one-sixth to one-thirtieth of the capacity is needed). Among the various techniques that cooperate to reach this objective i.e., reduce the amount of data to be transported, a very significant part comes from the observation that much information is repeated from frame to frame in a video signal. Therefore, by not re-transmitting portions of the picture which don't change between frames more compression is gained. Thus, MPEG breaks the video frames down into 8 x 8 regions called blocks. Four of these regions can be put together to create a 16 x 16 macroblock. Macroblocks that do not change are not re-encoded in subsequent frames. The macro-blocks are then grouped together into runs of macroblocks called slices. Typically the entire frame is covered by slices. The slice structure allows the receiver to re-synchronize at the beginning of a slice in case of data corruption, because each slice



begins with a unique header. Frames that can be predicted from previous frames are called P-frames. To avoid the propagation of errors and to allow periodic resynchronization, a complete frame which does not rely on information from other frames is transmitted approximately once every 12 frames, or 2-3 times a second. These stand-alone frames, called I-frames, are said to be intra-coded. There is also a third kind of frame which predicts pixel values from frames that occur both before and after it, these bi-directional frames are called B-frames.

Moreover, the prediction is even taken one step further by encoding motion vectors so that portions of a picture that move can be borrowed from other parts of previous frames of the video. Clearly, from this brief overview of MPEG frames, not all the transmitted frames of a video flow are of equal importance. If I-frames indeed deserve to receive the highest priority and should not be dropped, since they allow periodic resynchronization, decoders receiving MPEG video are devised so that they can easily accommodate occasionally the loss or the late transmission of other frame types without experiencing any noticeable degradation of the images to the casual viewer. This kind of compression scheme in which there is a hierarchy of the forwarded information, some pieces being essential while others are only used to improve the quality, is also used in the H.261 videoconferencing standard. Similarly, cellular phones conforming to the GSM (Global System for Mobile Communications) specifications promoted by the GSM Association (GSM Association, Avoca Court, Temple Road, Blackrock, Co. Dublin, Ireland), are using the same concept for the transmission of voice. That is, GSM encoding scheme manages to consider that some of the bits are more sensitive than others and are transmitted accordingly in frames and sub-frames, the less critical ones being discardable (especially, allowing a half-rate mode) however, without preventing phone conversations to go on normally even though it is occasionally at the expense of a poorer quality.

Thus, the concept of class of services does not fit well to flows of information carrying e.g., video and generally

other type of real-time compressed information, that must be granted high priority however, in which subflows of less importance can be defined, which do not really deserve the same level of priority. Then, in case of congestion those subflows  
5 are exhausting all the available bandwidth at a port of a network routing or switching device while it could better be attributed to other flows which, although they have a lower priority, are deemed to gain precedence when access to a port is completely denied.

## 10        **Object of the invention**

Therefore, it is a broad object of the invention to remedy the shortcomings of the prior art, as noted here above thus, enabling the possibility of processing differently, within priority flows, subflows that are not essential or are of less  
15 importance.

It is another object of the invention to permit that subflows be possibly discarded freeing some bandwidth to allow lower priority flows to be serviced too in case of congestion.

Further objects, features and advantages of the present  
20 invention will become apparent to the ones skilled in the art upon examination of the following description in reference to the accompanying drawings. It is intended that any additional advantages be incorporated herein.

### Summary of the Invention

A method and a system for allowing a selective discarding of data packets in a communications network, enforcing classes of service, are disclosed. The classes of service, organized from a lowest to a highest priority, are handling priority flows of data packets. The invention first assumes that subflows can be defined in the priority flows. The data packets, which belong to the subflows in the priority flows, are marked. When a traffic congestion window is detected the selective discarding of the invention is enabled. It consists in selecting, starting from the one of lowest priority, a contiguous set of subflows which are then discarded. While the traffic congestion lasts the selection of subflows is kept reassessed and the corresponding selected subflows continue to be discarded. When traffic congestion window ends the selective discarding is disabled and checking for the occurrence of a next traffic congestion window resume.

Therefore, the invention allows to process differently, within priority flows, subflows that are not essential or are of less importance, so that they are becoming eligible to discarding, freeing some bandwidth to allow lower priority flows to be serviced too in case of congestion.

### Brief Description of the Drawings

- Figure 1** shows a data communications network, or part of, comprised of three nodes.
- Figure 2** is a closer view of a node where invention may be exercised.
- Figure 3** describes a data packet housing in its header the priority information necessary to carry out the invention.
- Figure 4** shows the counters and their associated thresholds needed to implement the selective discarding of subflows.
- Figure 5** is a detailed description of the method of the invention.

## Detailed Description of the Preferred Embodiment

**Figure 1** illustrates the context where the invention preferably applies hence, showing a communication network (or part of) [100] comprising e.g., three nodes [110]. Nodes are aimed at steering data exchanged between endpoints [120] and [130] where they are generated, utilized and/or processed. Nodes are interconnected through communication lines as [140]. Depending upon technologies and communication protocols in use data are transported over permanent or temporary connections pre-established between endpoints [150] thus, are switched at each node or are rather routed from node to node based on the endpoint e.g., [120] and [130] destination addresses. Irrespective of the protocols and technologies employed nodes have ports as [160] through which in and out data are flowing.

**Figure 2** is a closer view of a network node as shown in figure 1 and implemented e.g., under the form of a switch [200]. Switch is comprised of a switch fabric [210] aimed at switching data packets like [220] and [230] through bi-directional ports and port adapters such as [240] and [250], including an egress buffer [252], further discussed in figure 4, and aimed at temporarily storing the flow of packets received from the switch fabric [210]. The type of problem that invention helps to manage is more likely to occur when an output port such as [251] is receiving much traffic, possibly simultaneously from many other input ports like [241] so that it becomes congested or is near congestion and discarding of packets must be considered because internal buffering is going to be exhausted i.e., egress buffer [252] maximum capacity is near to be reached.

**Figure 3** shows a generic data packet [300]. Although format, size and structure of data packets are largely dependent on protocols and technologies in use they all comprise a header [310] and a data payload [320], this latter carrying the end-users information. As far as switches are concerned data packets are generally fixed-size data packets. This permits to build very efficient switch fabrics able of handling aggregate throughput commonly measured in tenths, if not in hundreds, of Gigabits/sec. Header [310] contains at least all the necessary information to steer data packets within the switch fabric, towards output ports, as shown in figure 2. Whatever protocol and data packet format are actually in use the invention assumes that header contains two types of information necessary to carry out the invention. First, a subfield [330] containing the indication of the priority class (CoS) data packet belongs to. Second, a bit to indicate if the data packet is eligible to discarding in its priority class thus, carrying the indication that data packet belongs to a subflow as previously discussed. How and when the above information is set or modified in the header of the data packet is beyond the scope of the invention which rather focuses on how it is taken advantage of priority classes and subflows to better accommodate congestion windows at network nodes. Broadly speaking, it is up to the device and those in charge of managing the network and nodes to decide about priority classes of flows. Thus, network management i.e., the people in charge along with the hardware and software pieces of equipment they have put in place to control a network, is entitled to set the priority bit field [330] of the data packets exchanged between end points. However, as far as subflows are concerned, it is only the software and/or hardware components that have participated to the encoding of the source of information e.g., the video or audio signal encoders, that are actually able to distinguish those of the data that are essential from those that are eligible to

discarding eventually resulting in the formatting of data packets having the best effort discard bit [340] asserted.

In the following figures, that illustrate the best embodiment of the invention in a very high speed switch, four  
5 classes of priority are used namely: 0, 1, 2 and 3 [331]. Priority 0 is the highest priority. In this example header priority field [330] is thus a 2-bit field.

**Figure 4** shows the five counters, associated to each output port e.g., to a Nth port [400], and that are necessary  
10 to carry out the method of the invention. A main counter [405] takes care of the whole traffic destined to the output port it is attached to i.e., [400] in this example. Therefore, it is incremented [412] for each packet [410] destined to that port, irrespective of its priority class, and filling an egress  
15 buffer [402] aimed at temporarily storing the packets before they are actually leaving the current switching node. A second counter [415] counts all incoming packets for that port but those lowest priority (i.e., priority 3 in this example [416]) packets which have the 'best effort discard bit' asserted. In  
20 other words, second counter [415] includes the whole traffic of priorities 0, 1 and 2 plus the guaranteed traffic of priority 3, i.e., the packets with their 'best effort discard bit' unasserted, only excluding the best effort traffic of the  
lowest priority class i.e., the packets with their 'best  
25 effort discard bit' asserted.

Similarly:

- a third counter [420] counts all incoming packets except the best effort packets of priority 3 and 2 [421].
- a fourth counter [425] counts all incoming packets except  
30 the best effort packets of priority 3, 2 and 1 [426].

• finally, a fifth counter [430] counts all incoming packets except the best effort packets of priority 3, 2, 1 and 0 [431] hence, counts only those packets corresponding to the aggregate guaranteed bandwidth traffic of all priority classes.

Second to fifth here above counters are referred to as 'subflow delta counters' in the rest of this description. More specifically, second counter [415] is 'subflow delta counter (3)', third counter [420] is 'subflow delta counter (2-3)', forth counter [425] is 'subflow delta counter (1-3)' and fifth counter [435] is 'subflow delta counter (0-3)'. All here above counters i.e., the main counter and the subflow delta counters, are decremented [435] at a pace corresponding to the actual throughput of the output port [400] that is, are decremented whenever a packet is actually removed from the egress buffer [402] so as to be forwarded to a next switching node or to be processed in an end point as [130] in figure 1. Therefore, at any given moment, main counter [405] represents the actual filling of the egress buffer [402] while the other four counters are a measure of what would have this filling be if the corresponding priority subflows would have indeed been discarded. In other words, the set of five counters is a picture of the current history of the traffic, for that output port, from which switch fabric may derive its decision of what subflows have to be discarded should congestion occur.

Hence, associated to the main counter [405] there is an Enable Discard Threshold (EDT) [440] which, when reached or exceeded, enables the best effort discard mechanism per the invention. Obviously, EDT is set with respect to the actual size of the egress buffer, preserving a margin large enough to let the mechanism of the invention to take place well before indiscriminate discarding, that would result of the straight overflow of the egress buffer, may possibly occur. Also, associated to the value of the main counter [405], a Halt Discard Threshold (HDT) [445] is used to pause the discard process of the invention when packet count is equal or goes



below it. Conversely to HDT, which is aimed at marking the  
starting of a congestion window, EDT is ending it thus,  
disabling the selective discarding of the invention. Also,  
this second threshold (HDT) is set to a value low enough so  
5 that, no matter how big are the bursts of packets converging  
to the same output port, indiscriminate discarding of packets  
may possibly occur, just because corresponding egress buffer  
has suddenly overflowed. Finally, associated to the four  
subflow delta counters, there is a single common threshold  
10 used to select which one(s) of the subflow(s) have to be  
discarded. This is the Best Effort Priority Discard Threshold  
(BEPDT) [450]. BEPDT is set to a value, between EDT and HDT,  
so that the selective discarding of the invention better takes  
place, ideally obtaining that no discarding of packets,  
15 belonging to the guaranteed priority flows (i.e., those of the  
packets that do not have their best-effort discard set), ever  
occurs. Clearly, the three threshold values here above  
discussed, are highly dependent on the environment in which  
the invention is carried out, the objectives to reach and the  
20 applications that are taking advantage of it. Hence, it does  
not exist such a thing as ideal application independent values  
for those thresholds even though the objective to reach is  
clearly, as already mentioned here above, that discarding of  
packets, when necessary, only affects the less important  
25 subflows in each priority class. This may have to be obtained  
through a fine tuning of those thresholds.

In the particular example of figure 4 the subflow delta  
counter, of higher associated priority, with a value above  
BEPDT is the second one [420] from left i.e., subflow delta  
30 counter (2-3). Hence, it carries the indication that subflows  
of priority 3 and 2 must be discarded to ease congestion (in  
this example subflows 0 and 1, of higher priorities, are thus  
preserved). Discarding of those subflows is going to last as  
long as the picture of the counters stays alike. However, when  
35 one of the subflow delta counters is crossing BEPDT the situa-  
tion is immediately reassessed. In this particular example,

where subflows of priority 3 and 2 are discarded to ease congestion, it can reasonably be anticipated that value of subflow delta counter (2-3) soon goes below BEPDT in which case the mechanism of the invention calls for discarding only  
5 subflow of priority 3 since only the left subflow delta counter remains above BEPDT. The selective discarding of subflows lasts as long as there is at least one of the subflow delta counters above BEPDT.

**Figure 5** better describes the method of the invention  
10 which starts or resumes at step [500]. As long as the filling of the egress buffer stays below EDT selective discarding is in standby i.e., step [500] is looping on itself [505]. When EDT of main counter is crossed, detecting the beginning of a congestion window, the selective discarding per the invention  
15 is enabled [510]. Then, first step [520] consists in checking if the value of the subflow delta counter 0-3 is above BEPDT. If answer is positive then, subflow of priority 0 is discarded [525] which triggers the discarding of all the other subflows of lower priorities i.e., subflow 1 [535], subflow 2 [545] and  
20 subflow 3 [555]. It is worth noting here that although this description of the preferred embodiment of the invention utilizes four level of priorities it should be obvious to those skilled in the art that it could be practiced as well with fewer or more classes of services and still perform  
25 alike. Going back to step [520] if, however, the answer is negative then next checking step [530] consists in verifying if subflow delta counter 1-3 is above BEPDT. If answer is positive then, in a manner similar to what was just described for priority level 0, subflow 1 is discarded [535] which  
30 triggers in turn the discarding of subflows at lower priorities i.e., subflow 2 [545] and subflow 3 [555]. Therefore, the above process repeats for every remaining step [540] and [550], respectively corresponding to priority 2 and priority 3 thus triggering the discarding of subflow 2 and 3 at steps  
35 [545] and [555] or just to the discarding of subflow 3 [555] depending on the answers to the corresponding checkings.

However, if answers to checking steps [520] to [550] was constantly negative no discard of any subflow is necessary. Irrespective of the fact that discarding of subflows was necessary or not, method of the invention proceeds to step

5 [560] in order to check if main counter has gone below HDT, detecting the end of a congestion window. If answer is positive then, the selective discarding of the invention is disabled [570] and one returns [575] to first step [500] waiting for a next occurrence of a congestion window while, if

10 answer to step [560] is negative [565], one must resume at step [520] so that the subflow delta counters are reevaluated and acted on accordingly as previously described.



**Claims:**

What is claimed is:

1. A method for allowing a selective discarding of data packets in a communications network [100] enforcing a plurality of classes of service, said plurality of classes of service handling priority flows, said data packets belonging to said priority flows [330], said plurality of classes of service organized from a lowest to a highest priority [331], said method comprising the steps of:
  - 10     defining, in said priority flows, subflows;  
marking [340] those of said data packets of said priority flows that belong to said subflows;  
checking [500] for the occurrence of a traffic congestion window;
  - 15     upon entering said traffic congestion window:  
enabling said selective discarding [510], said step of enabling further comprising the steps of:
    - selecting [520, 530, 540, 550] a contiguous set of said subflows starting from said lowest priority [550];
    - 20     discarding [525, 535, 545, 555] said contiguous set of said subflows;
    - checking [560] for the ending of said traffic congestion window;
  - while remaining in said traffic congestion window:
    - 25     keep reassessing [565] said selection of said subflows;  
keep discarding said selection of said subflows;
  - upon leaving said traffic congestion window:
    - disabling said selective discarding [570];
    - resume [575] checking for the occurrence of a next said
    - 30     traffic congestion window.

2. The method according to claim 1 wherein said selective discarding is performed in an output port [251] of a node [110] of said communications network [100], said output port temporarily storing said data packets in an egress buffer [252] before said data packets can be forwarded.

3. The method according to any one of the previous claims wherein a main counter [405] monitors the filling of said egress buffer, said method further including the steps of:

associating, to said main counter, an enable discard threshold (EDT) [440] wherein said EDT corresponds to an upper value of filling of said egress buffer;

associating, to said main counter, a halt discard threshold (HDT) [445] wherein said HDT corresponds to a lower value of filling of said egress buffer;

4. The method according to any one of the previous claims wherein determining said traffic congestion window includes:

entering said traffic congestion window when value of said main counter is crossing upwards [500] said EDT;

leaving said traffic congestion window when value of said main counter is crossing downwards [560] said HDT.

5. The method according to any one of the previous claims wherein a subflow delta counter [415, 420, 425, 430] is attached to each of said plurality of classes of service, said method further including the steps of:

5 for all said subflow delta counters:

associating a best effort priority discard threshold (BEPDT) [450] wherein said BEPDT corresponds to an intermediate value of filling of said egress buffer [402];

for each of said subflow delta counters:

10 counting the number of said data packets currently filling said egress buffer however, excluding from count those of said data packets, belonging to said subflows, of priority equal to or less than the priority attached to said subflow delta counters [416, 421, 426, 431].

15 6. The method according to any one of the previous claims wherein the step of selecting a contiguous set of said subflows includes the steps of:

20 picking out, among said subflow delta counters, the one of higher priority having a value larger [420] than said BEPDT;

electing to discard said subflows of priorities ranging from said lowest priority up to the attached priority of picked out said subflow delta counter [421].

25 7. The method according to any one of the previous claims wherein said node is a fixed-size cell switch.

8. The method according to any one of the previous claims wherein a said priority flow is a MPEG video flow comprising a subflow formed from P-frames and B-frames.

9. A system, in particular a system implementing a selective discarding of data packets, comprising means adapted for carrying out the method according to any one of the previous claims.

- 5 10. A computer-like readable medium comprising instructions for carrying out the method according to any one of the claims 1 to 8.



**SYSTEM AND METHOD FOR ENABLING  
BEST-EFFORT SERVICE  
OF SUBFLOW PRIORITY FLOWS**

**Abstract**

5        A selective discarding of data packets in a communica-  
tions network, enforcing classes of service organized from a  
lowest to a highest priority, is disclosed. The classes of  
service, are handling priority flows of data packets. It is  
assumed that subflows that are not essential or are of less  
10 importance can be defined in the priority flows. Then, data  
packets, which belong to the subflows in the priority flows,  
are marked and become eligible to discarding when a traffic  
congestion window is detected. While traffic congestion lasts  
the selection of subflows to be discarded is kept reassessed.  
15 When traffic congestion window ends the selective discarding  
is disabled and checking for the occurrence of a next traffic  
congestion window resume. Therefore, the invention allows to  
process differently subflows of priority flows so that they  
may be discarded, freeing some bandwidth to allow lower prior-  
20 ity flows to be serviced too in case of congestion.

**Figure 4.**



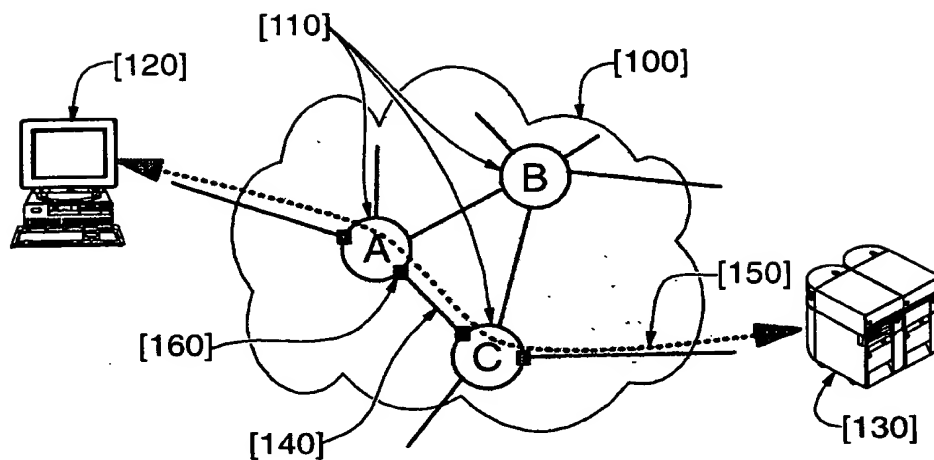


Figure 1

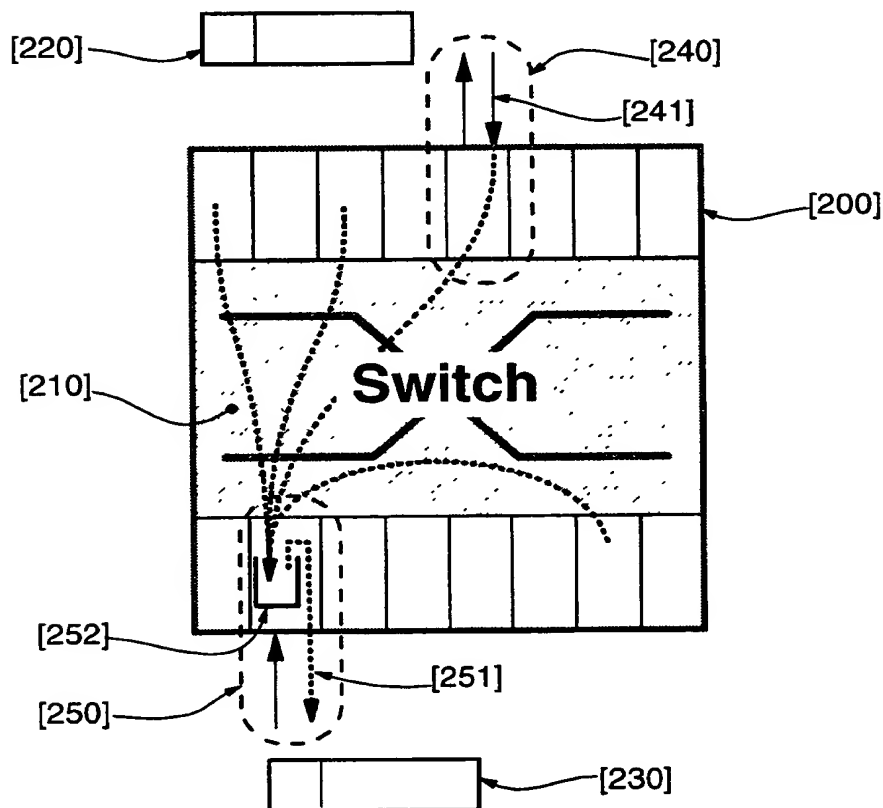


Figure 2

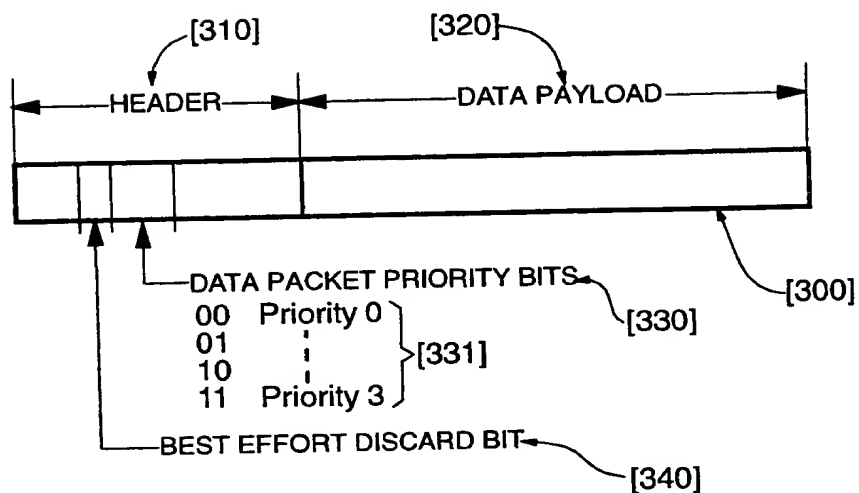


Figure 3

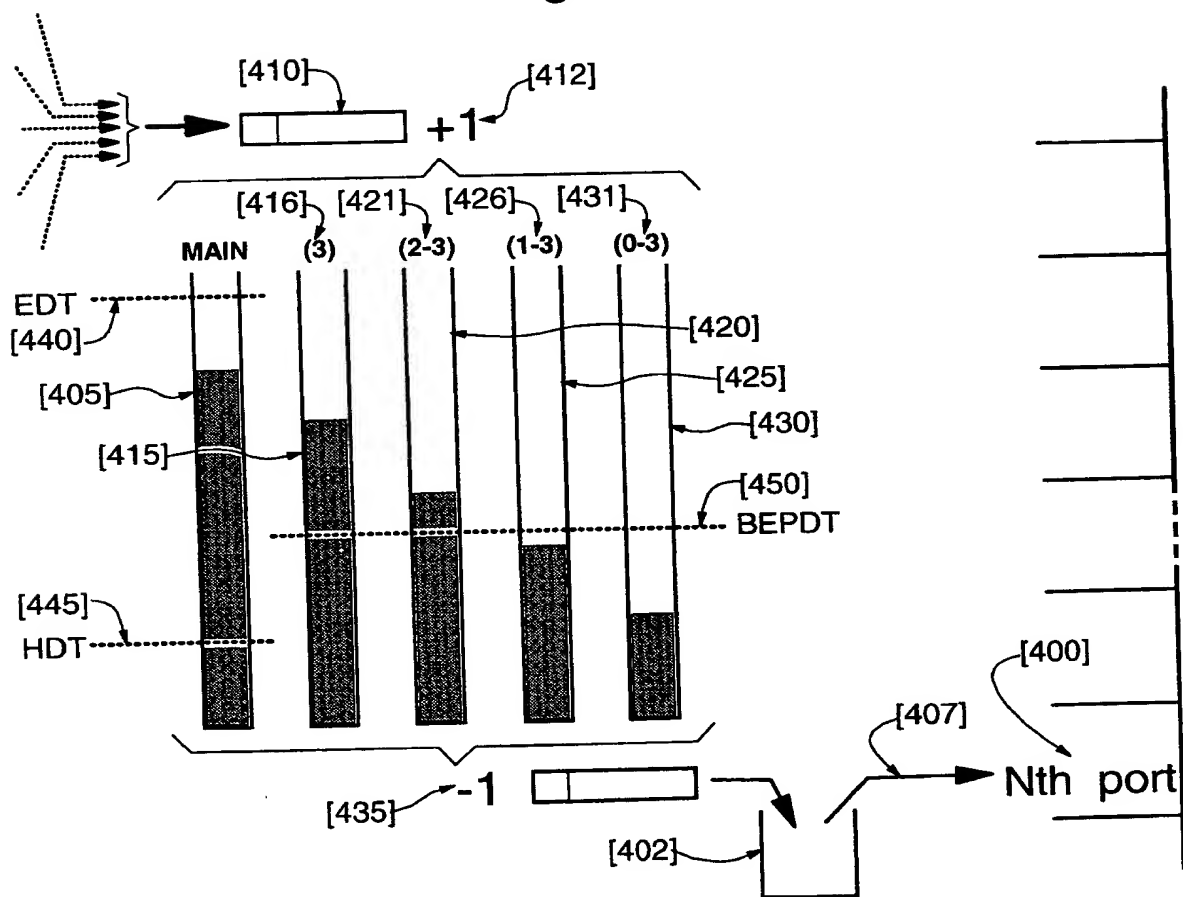


Figure 4

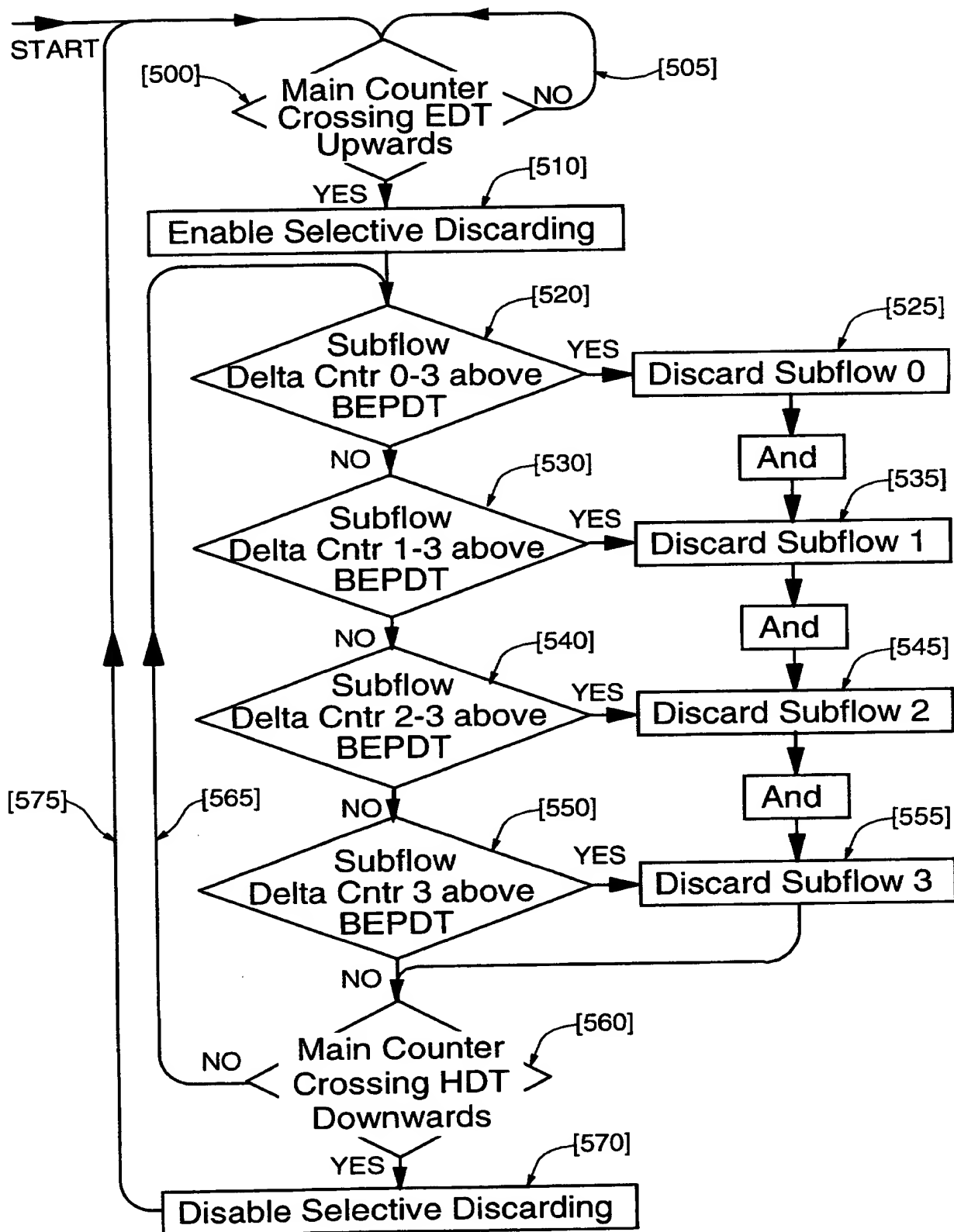


Figure 5

**This Page Blank (uspto)**